A Framework for Assessment Design in the Era of Generative AI: The Case of Take-Home Assignment in Software-related Courses

new Lyza FELIPEa\*, Quang Ngoc TRANb, Thanh Chi PHAMc, Thanh Ngoc NGUYENd & Linh Duc TRANe

a,b,c,d,e*School of Science, Engineering & Technology, RMIT University, Vietnam*

\*anna.felipe@rmit.edu.vn

**Abstract:** The use of generative AI in education has brought about various challenges, especially in the area of assessment design, particularly concerning academic integrity. Plagiarism is a significant concern as students can easily use AI-generated content to cheat in assignments, compromising the credibility and reliability of the assessment results. This paper addresses this issue by proposing a framework for designing take-home assignments in software-related courses incorporating generative AI and promoting academic integrity. The framework provides guidelines for designing assignments that prevent plagiarism and ensure that the assessment tasks align with the course learning objectives. It offers a practical and flexible approach to integrating generative AI into assessments while mitigating the risks associated with academic misconduct.

**Keywords:** AI Education, Bloom Taxonomy, Computing Education, Generative AI, Assessment Design

1. Introduction

In recent years, the field of education has witnessed a significant shift toward the integration of technology, particularly in the assessment design (Almond et al., 2002; Gorin, 2006; Van den Berg et al., 2006; Villarroel et al., 2018). One area where this integration has gained significant attention is the use of generative AI in assessment tasks (Suzanne Fergus et al., 2023; Geerling et al., 2023). Generative AI refers to computer algorithms that can generate new content or responses based on a set of input data (Will Yeadon et al., 2023). The integration of generative AI in assessment tasks provides various benefits such as efficient grading, increased objectivity, and reduced bias in the assessment (Benuyenah, 2023). However, the use of generative AI in assessment tasks poses several challenges, especially in the area of academic integrity (Kooli, 2023).

Academic integrity is a critical aspect of education that ensures that students maintain high ethical standards in their academic work (Emenike & Emenike, 2023; Suzanne Fergus et al., 2023). Plagiarism, the act of presenting someone else's work as one's own, is a significant concern in academic integrity, as it undermines the credibility and reliability of assessment results (Bretag et al., 2019). The use of generative AI in assessment design can facilitate plagiarism, as students can easily use AI-generated content to cheat on assignments (Crawford et al., 2023). This issue poses a significant challenge to educators, who must ensure that assessment tasks are designed to promote academic integrity while also taking advantage of the benefits of the generative AI (Crawford et al., 2023).

This paper addresses the challenges associated with the integration of generative AI in assessment design, specifically in the context of take-home assignments in software-related courses (Gilson et al., 2023). The researchers propose a framework that provides guidelines for designing assessment tasks that incorporate generative AI, while also promoting academic integrity (Mike Perkins, 2023). The proposed framework offers a practical and flexible approach for educators seeking to integrate generative AI into their assessment design process (Suzanne Fergus et al., 2023).

The paper is organized as follows: Section 2 provides a brief overview of the related work in the area of generative AI and assessment design. Section 3 describes our research approach for designing take-home assignments that incorporate generative AI. Section 4 presents findings that demonstrate the application of the proposed framework in the context of software-related courses. Finally, Section 5 concludes the paper and provides directions for future research in the area of generative AI and assessment design.

1. Related Research
   * 1. Bloom taxonomy

Bloom's Taxonomy is a well-known framework that has been used for several decades to describe the various levels of cognitive skills required for learning. Developed by Benjamin Bloom in the 1950s, the taxonomy identifies six levels of cognitive skills, which are arranged in a hierarchy, starting from the lowest level of recall or memory to the highest level of creation or synthesis. The six levels of Bloom's Taxonomy include: remembering, understanding, applying, analyzing, evaluating, and creating. The taxonomy has become a popular tool for educators worldwide to design learning objectives, the curriculum, and assessments that focus on developing students' cognitive skills (Forehand, 2010). It has been widely used in various fields, including education, psychology, and business, to design effective instructional strategies that enhance learning outcomes.

The framework has undergone several revisions over the years, with the most recent update released in 2001 by a team of educational psychologists led by Lorin Anderson. The revised taxonomy includes updated descriptions of each cognitive skill level, reflecting the current understanding of learning and teaching. Table 1 presents the comparison of generic questions of Bloom’s Taxonomy and Computer-Related Topics Questions of Bloom’s Taxonomy. Both share the same cognitive levels of Remembering, Understanding, Applying, Analyzing, Evaluating, and Creating. However, the content and context of the questions differ between the two sets. The questions in generic questions are general in nature and can be applied to a wide range of subjects or topics while computer-related topics questions focus on concepts, principles, and skills within the field of computer science and technology.

Table 1: Comparison of Generic Questions and Computer-Related Topics Questions

|  |
| --- |
| Computer-related topics questions of Bloom Taxonomy |
| 1. Remembering:    * What is the difference between a compiler and an interpreter?    * What is the syntax for creating a new object in Java?    * What is the purpose of a firewall in network security? 2. Understanding:    * How do programming languages like Java and Python differ in their approach to memory management?    * What is the role of an operating system in managing computer resources?    * How does the client-server model work in web development? 3. Applying:    * Can you write a program in Python that takes user input and performs a specific task?    * How would you design a database schema to store information about users in a web application?    * Can you troubleshoot and solve a common networking issue? 4. Analyzing:    * How would you identify and fix a performance bottleneck in a software application?    * How would you analyze the security risks associated with a particular system or application?    * Can you break down the steps involved in designing and implementing a complex algorithm? 5. Evaluation:    * How would you assess the effectiveness of a particular software development methodology?    * Can you critique the design of a user interface and suggest improvement?    * How would you evaluate the ethical implications of using artificial intelligence in a particular application? 6. Creating:    * Can you design and implement a new software application that solves a real-world problem?    * How would you develop a new feature or functionality for an existing software product?    * Can you create a prototype of a new technology that has not yet been developed? |

* + 1. Assessment design

Assessment design is a critical aspect of higher education, as it provides important information on student learning and informs ongoing improvement efforts (S. Fergus et al., 2023; W. Yeadon et al., 2023). One theme that emerges in the literature is the importance of aligning assessments with learning outcomes. (Chickering & Gamson, 1987) argue that assessment should be closely tied to course objectives, providing evidence that students are meeting the intended goals. This requires careful attention to the development of clear and specific learning outcomes, as well as the design of assessments that measure those outcomes (M. Perkins, 2023).

Another The use of generative AI in education has brought about various challenges, especially in the area of assessment design, particularly concerning academic integrity. Plagiarism is a significant concern as students can easily use AI-generated content to cheat in assignments, compromising the credibility and reliability of the assessment results. This paper addresses this issue by proposing a framework for designing take-home assignments in software-related courses incorporating generative AI and promoting academic integrity. The framework provides guidelines for designing assignments that prevent plagiarism and ensure that the assessment tasks align with the course learning objectives. It offers a practical and flexible approach to integrating generative AI into assessments while mitigating the risks associated with academic misconduct. consideration in assessment design is the use of rubrics. Many scholars emphasize the importance of rubrics in ensuring consistency and fairness in grading, as well as providing valuable feedback to students. (Brookhart, 2013) notes that rubrics can help students understand the expectations for their work and can encourage deeper engagement with the material. The literature also highlights the importance of incorporating a variety of assessment types in higher education (Mellar et al., 2018). This includes both formative assessments, designed to provide ongoing feedback and support to students, and summative assessments, used to evaluate learning at the end of a course or program. Brookhart (2013) notes that different assessment types may be more appropriate for different learning outcomes and that instructors should carefully consider which types will provide the most valid and reliable data.

1. references

Almond, R., Steinberg, L., & Mislevy, R. (2002). Enhancing the design and delivery of assessment systems: A four-process architecture. *The Journal of Technology, Learning and Assessment*, *1*(5).

Benuyenah, V. (2023). Commentary: ChatGPT use in higher education assessment: Prospects and epistemic threats. *Journal of Research in Innovative Teaching & Learning*, *16*(1), 134-135.

Bretag, T., Harper, R., Burton, M., Ellis, C., Newton, P., Rozenberg, P., Saddiqui, S., & van Haeringen, K. (2019). Contract cheating: A survey of Australian university students. *Studies in Higher education*, *44*(11), 1837-1856.

Brookhart, S. M. (2013). *How to create and use rubrics for formative assessment and grading*. Ascd.

Chickering, A. W., & Gamson, Z. F. (1987). Seven principles for good practice in undergraduate education. *AAHE bulletin*, *3*, 7.

Crawford, J., Cowling, M., & Allen, K.-A. (2023). Leadership is needed for ethical ChatGPT: Character, assessment, and learning using artificial intelligence (AI). *Journal of University Teaching & Learning Practice*, *20*(3), 02.

Emenike, M. E., & Emenike, B. U. (2023). Was This Title Generated by ChatGPT? Considerations for Artificial Intelligence Text-Generation Software Programs for Chemists and Chemistry Educators. *Journal of Chemical Education*, *100*(4), 1413-1418.

Fergus, S., Botha, M., & Ostovar, M. (2023). Evaluating academic answers generated using chatgpt. *Journal of Chemical Education*.

Fergus, S., Botha, M., & Ostovar, M. (2023). Evaluating Academic Answers Generated Using ChatGPT [Article]. *Journal of Chemical Education*, *100*(4), 1672-1675. <https://doi.org/10.1021/acs.jchemed.3c00087>

Forehand, M. (2010). Bloom’s taxonomy. *Emerging perspectives on learning, teaching, and technology*, *41*(4), 47-56.

Geerling, W., Mateer, G. D., Wooten, J., & Damodaran, N. (2023). ChatGPT has aced the test of understanding in college economics: Now what? *The American Economist*, 05694345231169654.

Gilson, A., Safranek, C. W., Huang, T., Socrates, V., Chi, L., Taylor, R. A., & Chartash, D. (2023). How does CHATGPT perform on the United States Medical Licensing Examination? the implications of large language models for medical education and knowledge assessment. *JMIR Medical Education*, *9*(1), e45312.

Gorin, J. S. (2006). Test design with cognition in mind. *Educational measurement: Issues and practice*, *25*(4), 21-35.

Kooli, C. (2023). Chatbots in education and research: a critical examination of ethical implications and solutions. *Sustainability*, *15*(7), 5614.

Mellar, H., Peytcheva-Forsyth, R., Kocdar, S., Karadeniz, A., & Yovkova, B. (2018). Addressing cheating in e-assessment using student authentication and authorship checking systems: Teachers' perspectives [Article]. *International Journal for Educational Integrity*, *14*(1), Article 2. <https://doi.org/10.1007/s40979-018-0025-x>

Perkins, M. (2023). Academic Integrity considerations of AI Large Language Models in the post-pandemic era: ChatGPT and beyond. *Journal of University Teaching & Learning Practice*, *20*(2), 07.

Perkins, M. (2023). Academic Integrity considerations of AI Large Language Models in the post-pandemic era: ChatGPT and beyond [Article]. *Journal of University Teaching and Learning Practice*, *20*(2), Article 7. <https://doi.org/10.53761/1.20.02.07>

Van den Berg, I., Admiraal, W., & Pilot, A. (2006). Design principles and outcomes of peer assessment in higher education. *Studies in Higher education*, *31*(03), 341-356.

Villarroel, V., Bloxham, S., Bruna, D., Bruna, C., & Herrera-Seda, C. (2018). Authentic assessment: creating a blueprint for course design. *Assessment & Evaluation in Higher Education*, *43*(5), 840-854.

Yeadon, W., Inyang, O.-O., Mizouri, A., Peach, A., & Testrow, C. P. (2023). The death of the short-form physics essay in the coming AI revolution. *Physics Education*, *58*(3), 035027.

Yeadon, W., Inyang, O. O., Mizouri, A., Peach, A., & Testrow, C. P. (2023). The death of the short-form physics essay in the coming AI revolution [Article]. *Physics Education*, *58*(3), Article 035027. <https://doi.org/10.1088/1361-6552/acc5cf>